

## Description

# STRUCTURES AND METHODS FOR MANUFACTURING STRAINED P-TYPE MOSFET WITH GRADED EMBEDDED SILICON-GERMANIUM SOURCE-DRAIN EXTENSION

### BACKGROUND OF INVENTION

- [0001] The present invention relates generally to the field of semiconductor manufacturing and of integrated circuit fabrication. More particularly, the present invention relates to complementary metal oxide semiconductor (CMOS) field effect transistors (FET) with strained silicon for high performance and a method for manufacturing the same.
- [0002] Since it has become increasingly difficult to improve metal-oxide-semiconductor-field-effect transistor (MOSFET) performance through continued scaling, methods enhancing performance of MOSFETs without scaling

have become important. Strained silicon (Si) has shown enhancement of mobility for both electrons and holes. Therefore, in contemporary CMOS technology, there is significant interest and work in the use of strained material for the FET channel.

[0003] In one approach, silicon-germanium alloy (SiGe) is used to form a surface channel strained Si/relaxed SiGe n-type MOSFET (NMOSFET or NFET). In that approach, biaxial tensile strain is induced in a very thin epitaxial Si layer. The tensile strain reduces the effective mass of the electrons, which results in enhancement of electron mobility.

[0004] In the case of a p-type MOSFET (PMOSFET or PFET), the germanium (Ge) concentration must be greater than about 30% in order to have an effective increase in hole mobility.

[0005] This approach has the following drawbacks: 1) The strained silicon is grown on relaxed SiGe, and therefore it is difficult to control the leakage of devices.

[0006] 2. The requirement for enhanced performance of more than 30% germanium concentration further increases the defect density.

[0007] 3. The high diffusivity of dopant, such as arsenic and phosphorus, in SiGe makes it difficult to form shallow p-n junctions. For sub-micron or deep-sub-micron MOSFETs,

shallow junctions are required to scale devices.

[0008] Thus, the art has a need for a method of making circuits having the benefits of strained silicon without the low yield characteristic of the prior art.

## **SUMMARY OF INVENTION**

[0009] The invention relates to a method of forming a PMOSFET having compressive stress on the channel of PMOSFETs, thereby improving hole mobility.

[0010] A feature of the invention is the introduction of compressive stress in selected areas of an integrated circuit by introduction of germanium into the silicon lattice in the areas that will become the source and drain.

[0011] A feature of the invention is the use of a PMOSFET with graded germanium doped source/drain and/or extension for strain development provided by an epitaxial layer of SiGe. Since graded germanium profile does not have a sharp interface (where the dislocations are generated) between SiGe and Si, dislocations are not easily produced. This can reduce leakage of devices caused by dislocations. Graded germanium profiles also provide a method to optimize stress in order to enhance device performance.

[0012] Another feature of the invention is full or partial vertical penetration by germanium of the SOI layer to form areas

of SiGe within the SOI layer.

[0013] Another feature of the invention is a choice between diffusion driven by an anneal and/or diffusion driven by an oxidation process.

#### **BRIEF DESCRIPTION OF DRAWINGS**

[0014] Figure 1 illustrates a structure formed with one version of the invention. A SOI PMOSFET structure in which the graded SiGe source/drain (SD) reaches to the buried insulator.

[0015] Figure 2 illustrates a structure formed with a second version of the invention. A SOI PMOSFET structure in which the graded SiGe SD does not reach to the buried insulator.

[0016] Figure 3 illustrates a structure formed with a third version of the invention. A bulk PMOSFET structure with graded SiGe SD.

[0017] Figure 4 illustrates an initial step in the formation of the structure of Figure 1.

[0018] Figure 5 illustrates patterning the gate stack.

[0019] Figure 6 illustrates forming the epitaxial dopant layer.

[0020] Figure 7 illustrates the structure after diffusing the germanium into the silicon SOI layer.

[0021] Figure 8 illustrates a structure that is an alternative to the

structure of Figure 7.

[0022] Figure 9 illustrates a following step to the step of Figure 8.

#### **DETAILED DESCRIPTION**

[0023] Figure 1 shows in cross section a simplified field effect transistor 100 formed in a portion of a silicon wafer 10 having a buried insulator layer 20 and a device layer of silicon or semiconductor (or an SOI layer) 30. Transistor gate 110 and sidewalls 105 are positioned above gate oxide 103 that defines the lateral extent of body 30. Source and drain 40 are formed from portions of the SOI layer 30 that have been doped with germanium in a previous step. The germanium has diffused downward toward the insulator layer 30 and inward to the center of the channel under the gate. This has imposed compressive stress in the horizontal direction (parallel to the SOI surface) in the SOI and tensile stress in vertical direction (normal to the SOI surface) in the center of the device. These components complete a field effect transistor that illustratively has an n-type body that forms a channel for p-type field effect transistor (PMOSFET).

[0024] A PFET is shown for illustration. Typically, in current technology, the circuit will use CMOS technology, including

both NMOSFETs and PMOSFETs.

[0025] The wafer is illustratively a commercially available wafer produced by the SIMOX (Separation by Implantation of Oxygen) process. Other methods of producing wafers may also be used.

[0026] Referring back to Figure 1, the source and drain 40 on either side of the gate have a gradation of germanium having a maximum value at the top and reducing in magnitude toward the BOX (buried oxide) layer 20. This is caused by the diffusion of germanium into the silicon SOI layer from an upper layer of SiGe or germanium. Implantation of germanium is a possible alternative, but for most applications is unsatisfactory because it would take too long to deliver the required dose. In addition, implantation of a typical dose of germanium would cause extensive crystal damage that would be difficult to remove.

[0027] An advantage of the use of an epitaxial step for the deposition of the SiGe layer is that the epitaxial deposition is selective and deposits the SiGe only on exposed silicon. This reduces the amount of cleanup that would be required otherwise for the case of depositing SiGe or Ge everywhere over the device (e.g. gate, spacer, and SD).

[0028] The diffusion process may be carried out long enough to

spread the germanium concentration fairly evenly or it may be limited in time so that there is a significant concentration gradient from top to bottom of the SOI layer.

[0029] The presence of the germanium in the silicon lattice creates a compressive stress in the source and drain extending from left to right in the Figures. This stress, in turn, produces a compressive stress in the horizontal direction (parallel to the SOI surface) in the SOI and tensile stress in the vertical direction (normal to the SOI surface) in the channel of the device 100.

[0030] Both compressive stress in horizontal direction in the SOI and tensile stress in the vertical direction in the channel of the device produces increased hole mobility that increases PMOSFET performance.

[0031] Preferably, the germanium concentration is graded from top to bottom i.e. the diffusion is limited in time so that the germanium does not spread uniformly throughout layer 30, especially does not spread into the channel area of the device. The concentration has a maximum value at the top and declines to a dopant depth that is less than the thickness of the SOI layer. This gradation produces fewer dislocations in the crystal structure than a sharp profile would.



[0032] Referring now to Figure 4, the starting material for the practice of the invention is illustrated, in which substrate 10 supports BOX 20 having SOI layer 30 positioned above it.

[0033] In contemporary technology, the thickness of layer 30 may be in the range of 5 100 nm. Such thin layers are difficult to process with conventional methods, and it is a beneficial aspect of the invention that thin SOI layers are not a problem.

[0034] For a PMOSFET, boron is usually used to dope the SD area and arsenic to dope the channel area. Boron diffusion is slower in SiGe than in Si. Arsenic diffusion is slower in Si than in SiGe. The resulting structure can help to form shallow SD and an extension p-n junction and a sharp halo profile of the PMOSFET.

[0035] In Figure 4, a gate oxide layer 103, nominally 0.5 2 nm thick has been grown on the surface of layer 30. An alternative gate insulator layer, such as nitride, oxy-nitride or high-K materials might also be used.

[0036] A gate layer of polycrystalline silicon (poly) 110 has been deposited on oxide layer 30, with a hardmask layer of nitride 107 on the top.

[0037] Figure 5 shows the result of depositing, exposing and



curing a layer of photoresist 108, then etching a gate stack to be used in the FET to be formed. Illustratively, a reactive ion etch (RIE) is used because of its directional properties. Conventional etch materials are used as required to etch the various materials cleanly, stopping on the top surface of SOI layer 30.

[0038] Figure 6 shows the result of forming a thin conformal layer of nitride, illustratively about 10nm thick, then etching the horizontal portions of the nitride to leave an encapsulating layer 117 that protects the sides of the gate stack during subsequent processing and also defines a displacement distance between the transistor body below gate oxide 103 and the germanium diffusion that will follow. The portion of layer 117 on the top of gate 110 is the remainder of layer 107 after the step of etching the horizontal nitride surface.

[0039] A selective epitaxial growth process has also been performed, growing the epitaxial material 130 on exposed silicon surfaces.

[0040] Epitaxial layer 130 has a germanium concentration that is preferably greater than about 30% and a thickness such that sufficient germanium is available to diffuse into the SOI layer 30. Illustratively, the thickness of epitaxial layer

130 is about 15–30nm.

[0041] Figure 7 shows the result of performing an anneal step at 1000C for 10 minutes, which produces the desired diffusion of Germanium from dopant layer 130 into the SOI layer 30, with a maximum germanium concentration value at the top of layer 30, decreasing to a lower value at the bottom. The temperature and time of the anneal will be selected to produce the desired germanium gradient; in this case, having a significant germanium concentration all the way to the bottom surface of layer 30, forming SD 40.

[0042] The addition of germanium to the crystal structure of the silicon SOI layer will impose a compressive strain on the material in the source and drain 40 in Figure 1. That compressive strain will, in turn, exert compressive stress in the horizontal direction in the SOI and tensile stress in the vertical direction in the channel of the device 100.

[0043] Conventional later steps in the process, well known in the art, such as performing a halo implant, an extension implant, spacer formation, S/D implantation and metallization are referred to for purposes of the claims as "completing the transistor". This approach can be combined with a disposable spacer method to manufacture a high perfor-

mance MOSFET. That is removing the nitride spacer, re-oxidation of the poly gate as needed, halo and extension ion implantation, spacer formation, followed by SD ion implantation and SD anneal.

[0044] In summary, the process is: Start with a silicon SOI wafer; grow gate oxide (or equivalent); deposit poly gate layer (or deposit metal for metal gate device); deposit nitride mask layer; Deposit and pattern photoresist, RIE nitride, poly-Si (RIE metal for metal gate) and oxide; form thin nitride spacer (~ 10–30 nm); selective epitaxial SiGe (or germanium) on exposed silicon; high temperature anneal, diffusing germanium into Si device layer, grading germanium concentration vertically and horizontally, but preferably the germanium profiles do not enter the channel area; complete transistor including halo implant, extension implant, spacer formation, S/D implant, RTA, metallization.

[0045] Figure 8 shows a step in a second version of the invention in which the steps up to and including Figure 6 are the same. The resulting structure produced by this alternative method is the same as that in Figure 1. Figure 8 shows the growth of an oxide layer 135 on the top surface of SiGe layer 130. In this example, the oxide has consumed

the entire SiGe layer, turning it into oxide 135. Due to the snowplow effect of the oxidation in SiGe, Ge atoms in SiGe are driven into SOI layer. The oxidation approach is more efficient to drive in Ge into Si than annealing only. However, an oxidation method may produce defects that cause leakage of devices. Also it may be not easy to control the oxidation process at corners of the SiGe.

[0046] During the course of the thermal oxidation, the heat required has driven Germanium into SOI layer 30, in this case, extending all the way down to BOX layer 20 and forming graded SiGe SD 40.

[0047] Figure 9 illustrates the result of stripping oxide 135 in a wet etch, leaving a clean surface for subsequent processing.

[0048] The structure of Figure 9 may also be completed with the same conventional completion steps as discussed above in connection with Figure 7.

[0049] Figure 2 shows an alternative version of the invention, in which the Source and Drain 40 extend only partially through the SOI layer, leaving a portion of SOI layer 30 as lower silicon 35. The gate structure is the same in Figures 1 and 2.

[0050] The structure of Figure 2 has the advantage that the Ger-

manium is concentrated at the top of SOI layer 30, so that the stress is concentrated there also. For a surface channel transistor, the presence of strain lower down in the SOI layer is not of any substantial benefit, while the diffusion to a shallower depth will take a shorter time.

[0051] The structure shown in Fig. 1 is particularly important for ultra-thin SOI devices since it is very difficult to controllably etch SD area down without reaching the BOX. However, it is required (by the conventional method of embedded SiGe SD) to have Si left in the SD area in order to epitaxially grow SiGe in the SD area. Undesirable misfit dislocations can be generated when the thickness of strained SiGe exceeds the critical thickness for a given Ge fraction. This restricts the use of high Ge fractions in strained SiGe with large thickness. One can vary the structures between Fig. 1 and Fig. 2 to optimize the stress in the channel. For example, a high Ge fraction with a thin graded SiGe SD or a low Ge fraction with a thick graded SiGe SD.

[0052] Either of the two methods discussed above may be used to generate this structure, in which the diffusion of germanium into the SOI layer is controlled to stop before the entire layer has been filled with Germanium.

[0053] In forming a structure as shown in Figure 2, it is necessary to limit the exposure to heat, in order to limit the vertical extent of germanium diffusion.

[0054] Figure 3 shows another alternative version of the invention, in which the wafer is a bulk silicon wafer, so that the Source and Drain 40 extend only partially into the bulk silicon, leaving a portion of lower silicon 10. Another way to form the Fig. 3 structure is that of etching down Si in SD area followed by selective epi SiGe (as proposed by prior art). An anneal then forms graded SiGe in the SD area. This method has advantage of producing stronger stress in the channel with the epi interface away from the channel. Therefore, this can reduce the leakage caused by the defects at an epi interface.

[0055] The advantages and disadvantages of bulk versus SOI wafers are well known in the art and need not be repeated here. The designer of the integrated circuit will have to make a trade off in order to select bulk or SOI wafers.

[0056] This version of the invention may also make use of the annealing or the oxidation method of diffusion.

[0057] While the invention has been described in terms of a single preferred embodiment, those skilled in the art will recognize that the invention can be practiced in various

versions within the spirit and scope of the following claims.